

SPEECH RECOGNITION SYSTEM AND RECORDING MEDIUM

Publication number: JP2002041085

Publication date: 2002-02-08

Inventor: KITAOKA NORIHIDE; ONO HIROSHI

Applicant: DENSO CORP

Classification:

- international: G01C21/00; G06F3/16; G08G1/0969; G10L13/00; G10L15/00; G10L15/06; G10L15/08; G10L15/28; G10L15/18; G10L15/22; G01C21/00; G06F3/16; G08G1/0969; G10L13/00; G10L15/00; (IPC1-7); G10L15/28; G01C21/00; G06F3/16; G08G1/0969; G10L13/00; G10L15/00; G10L15/06

- european: G01C21/36; G10L15/08

Application number: JP20000220564 20000721

Priority number(s): JP20000220564 20000721

Also published as:



US6937982 (B2)

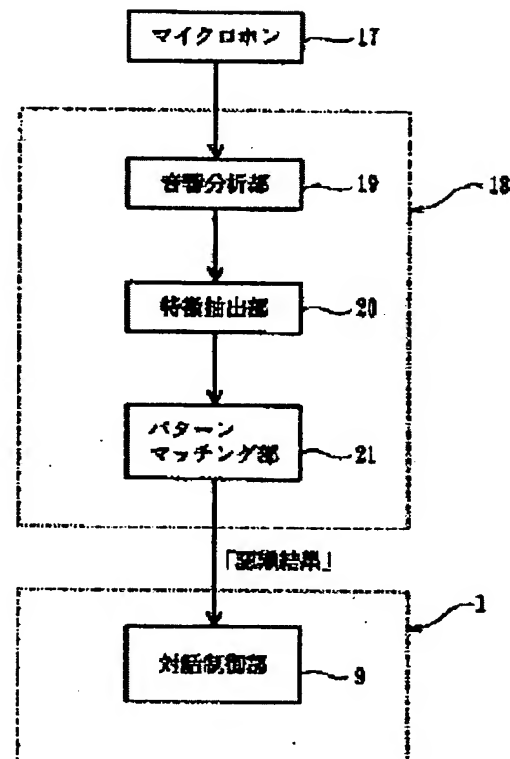
US2002010579 (A1)

DE10134908 (A1)

Report a data error here

Abstract of JP2002041085

PROBLEM TO BE SOLVED: To reduce the possibility that an input voice frequency signal is erroneously recognized as a word which causes an extraordinary working to peripheral equipment when the result of recognition of the input voice frequency signal is subjected to an operational control of the peripheral equipment. **SOLUTION:** A pattern matching element 21 executes pattern matching for the patterns between the time series data of LPC Cepstrum coefficient for the input voice frequency signal extracted in a characteristic-extracting element 20 and the standard pattern of a recognition-objective vocabulary which is preliminarily stored in the dictionary part. As the standard pattern of the recognition-objective vocabulary, not only the standard pattern corresponding to a specific word which possibly causes the state of an erroneous action in the case that it is given to the peripheral equipment as a voice command and to the second specific word which is adversative to the above specific word respectively but also a group of the standard patterns which correspond to the voice group similar to the specific word and the second specific word each, respectively, are prepared as a recognition-objective candidate group. The pattern matching element 21 outputs the second specific word as the result of the recognition when the standard pattern, which highly accords with the time series data of LPC cepstrum coefficient, exists in the recognition-objective candidate group.



(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3567864号

(P3567864)

(45) 発行日 平成16年9月22日(2004.9.22)

(24) 登録日 平成16年6月25日(2004.6.25)

(51) Int. Cl.⁷

F I

G 1 0 L 15/28

G 1 0 L 3/00 5 6 1 H

G 1 0 L 15/00

G 1 0 L 3/00 5 2 1 V

G 1 0 L 15/06

G 1 0 L 3/00 5 5 1 Q

請求項の数 7 (全 12 頁)

(21) 出願番号	特願2000-220564 (P2000-220564)	(73) 特許権者	000004260
(22) 出願日	平成12年7月21日(2000.7.21)		株式会社デンソー
(65) 公開番号	特開2002-41085 (P2002-41085A)		愛知県刈谷市昭和町1丁目1番地
(43) 公開日	平成14年2月8日(2002.2.8)	(74) 代理人	100096998
審査請求日	平成13年6月8日(2001.6.8)		弁理士 碓氷 裕彦
前置審査		(74) 代理人	100118197
			弁理士 加藤 大登
		(74) 代理人	100123191
			弁理士 伊藤 高順
		(72) 発明者	北岡 敦英
			愛知県刈谷市昭和町1丁目1番地 株式会
			社デンソー内
		(72) 発明者	大野 宏
			愛知県刈谷市昭和町1丁目1番地 株式会
			社デンソー内

最終頁に続く

(54) 【発明の名称】 音声認識装置及び記録媒体

(57) 【特許請求の範囲】

【請求項1】

話者により入力された音声信号を音響分析する音響分析手段と、
この音響分析手段での分析結果に基づいて入力音声信号の特徴パラメータを抽出する特徴
抽出手段と、

この特徴抽出手段で抽出された特徴パラメータと予め記憶されている認識対象語彙の比較
対象パターンとのパターンマッチングを行って一致度合いが高い単語を認識結果として出
力するパターンマッチング手段とを備え、

前記パターンマッチング手段による認識結果を外部機器の動作制御に供する音声認識装置
において、

前記認識対象語彙の比較対象パターンとして、前記外部機器の動作に悪影響を及ぼす可能
性がある特定の単語に対応した比較対象パターンの他に、その特定単語と類似した音声群
に対応した比較対象パターンの集団が認識対象候補集団として用意され、

前記パターンマッチング手段は、前記パターンマッチング時において、前記特徴抽出手段
で抽出された特徴パラメータとの一致度合いが高い比較対象パターンが前記認識対象候補
集団中に存在する場合に、前記特定単語を避けた単語を出力することを特徴とする音声認
識装置。

【請求項2】

請求項1記載の音声認識装置において

前記認識対象候補集団として、前記外部機器の機能低下を招く特定の単語と類似した音声

群に対応した比較対象パターンの集団が用意されることを特徴とする音声認識装置。

【請求項3】

話者により入力された音声信号を音響分析する音響分析手段と、
この音響分析手段での分析結果に基づいて入力音声信号の特徴パラメータを抽出する特徴抽出手段と、

この特徴抽出手段で抽出された特徴パラメータと予め記憶されている認識対象語彙の比較対象パターンとのパターンマッチングを行って一致度合いが高い単語を認識結果として出力するパターンマッチング手段とを備え、

前記パターンマッチング手段による認識結果を外部機器の動作制御に供する音声認識装置において、

前記認識対象語彙の比較対象パターンとして、前記外部機器を話者の意図と異なる状態で動作させる可能性がある特定の単語に対応した比較対象パターンの他に、その特定単語と類似した音声群に対応した比較対象パターンの集団が認識対象候補集団として用意され、前記パターンマッチング手段は、前記パターンマッチング時において、前記特徴抽出手段で抽出された特徴パラメータとの一致度合いが高い比較対象パターンが前記認識対象候補集団中に存在する場合に、前記特定単語を避けた単語を出力することを特徴とする音声認識装置。

【請求項4】

請求項3記載の音声認識装置において

前記認識対象候補集団として、前記外部機器の設定内容の変更を招く可能性がある特定の単語と類似した音声群に対応した比較対象パターンの集団が用意されることを特徴とする音声認識装置。

【請求項5】

前記パターンマッチング手段は、前記特定単語を避けた認識結果を出力する動作を、前記外部機器における所定の動作制御を許容するか否かの確認のために話者による音声入力を求める場面で行うように構成されていることを特徴とする請求項1ないし4の何れかに記載の音声認識装置。

【請求項6】

前記パターンマッチング手段による認識結果により動作制御される外部機器はナビゲーション装置であることを特徴とする請求項1ないし5の何れかに記載の音声認識装置。

【請求項7】

請求項1ないし6の何れかに記載された音声認識装置におけるパターンマッチング手段の機能をコンピュータシステムにおいて実現するためのプログラム及びデータを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、マンマシン・インタフェースとして使用される音声認識装置及びその音声認識装置の主要部の機能をコンピュータシステムにおいて実現するためのプログラム及びデータを記録した記録媒体に関する。

【0002】

【発明が解決しようとする課題】

マイクロホンを通じて入力された音声認識する音声認識装置にあっては、入力音声信号を分析した特徴パラメータと、辞書部に記憶されている認識対象語彙の比較対象パターンとを比較することにより、認識対象語彙中から比較結果が近い一つ（または複数）の単語を選択して出力する構成とされるものであるが、この場合、入力された音声認識対象語彙中に存在しない場合でも、何れかの類似単語を選択して出力する構成が一般的となっている。また、このような誤認識の他に、周囲雑音などに起因して入力音声信号とは異なる単語を出力してしまうという誤認識も不可避であるという一般的事情もある。このため、音声認識装置の出力により外部機器の動作制御を行う際には、上記のように入力音声信号

10

20

30

40

50

が誤認識された場合に、外部機器の動作に重大な悪影響を及ぼす可能性が出てくるものであり、その一例として、以下のような事例が考えられる。

【0003】

例えば、操作信号の音声入力並びにユーザ（車両運転者）側への質問、返答などの音声出力を行うボイスコントロール機能付きのカーナビゲーション装置において、周知のルート探索動作により設定されたルートを地図画面により案内させながら目的地へ向かって車両を走行させている途中で、ユーザが表示地図の縮小を意図して「しゅくしょう」というボイスコマンドを発声した場合に、装置側でこれを「じたく」と誤認識する可能性がある。この場合、仮に、カーナビゲーション装置が音声入力に応じて当該音声入力による指令内容を直ちに実行する構成であった場合には、上記のような誤認識が行われたときに、目的地までの設定ルートを破棄して自宅までのルート探索動作が行われることになり、結果的に、ユーザの意図に反して目的地までの設定ルートが破棄されてしまう。そこで、実際には、装置側では「自宅を目的地に設定します」、「よろしいですか?」と返答し、ユーザ側に確認を求める構成とされる。従って、ユーザ側では、上記確認を求められたときに「いいえ」と答えれば、装置側では自宅までのルート探索動作を行うことなく待機するようになり、改めて表示地図を縮小するためのボイスコマンドを入力すれば良い。しかし、ここでユーザ側が慌てて「しゅくしょう」などと発声した場合には、装置側でこれを「はい」と誤認識する可能性が生じ、このような場合には目的地までの設定ルートがユーザの意図に反して破棄されてしまうという問題点を引き起こす。

【0004】

要するに、音声認識装置の認識対象語彙中に、誤認識されることで、外部機器の動作に不都合が生じたり、外部機器がユーザの意図に反した動作を行う可能性がある単語が存在することがあり、従って、音声認識装置においては、上記のような誤認識に起因して外部機器が不用意な動作をする事態を未然に防止できる技術を採用することが望ましいものである。

【0005】

そこで、本発明の目的は、入力音声信号を認識結果を外部機器の動作制御に供する場合に、その入力音声信号を当該外部機器の異常動作を招く単語として誤認識する可能性を低くできて、外部機器が不用意な動作をする事態を未然に防止可能になる音声認識装置並びに上記誤認識の可能性を低くできる音声認識動作をコンピュータシステムにおいて実現可能なプログラム及びデータを記録した記録媒体を提供することにある。

【0006】

【課題を解決するための手段】

上記目的を達成するために請求項1に記載した手段を採用できる。この手段は、音響分析手段、特徴抽出手段及びパターンマッチング手段を備えており、話者による入力音声信号を音響分析すると共に、その分析結果に基づいて入力音声信号の特徴パラメータを抽出し、抽出された特徴パラメータと予め記憶されている認識対象語彙の比較対象パターンとのパターンマッチングを行って一致度合いが高い単語を認識結果とするものであり、その認識結果を外部機器の動作制御に供するようになる。この場合、認識対象語彙の比較対象パターンとして、前記外部機器の動作に悪影響を及ぼす可能性がある特定の単語に対応した比較対象パターンの他に、その特定単語と類似した音声群に対応した比較対象パターンの集団が認識対象候補集団として用意されており、パターンマッチング手段は、パターンマッチング時において、入力音声信号から抽出された特徴パラメータとの一致度合いが高い比較対象パターンが上記認識対象候補集団中に存在する場合に、前記特定単語を避けた単語を出力する。

【0007】

ここで、上記手段のように、認識対象語彙の比較対象パターン中に、外部機器の動作に悪影響を及ぼす可能性がある特定の単語に対応した比較対象パターンの他に、その特定単語と類似した音声群に対応した比較対象パターンの集団が用意された場合には、入力音声信号が特定単語であると認識される特徴空間が限定されることになる。このような状態は、特

10

20

30

40

50

定単語以外の単語が当該特定単語と認識される可能性が相対的に低くなった状態に相当するようになり、このときには、当該特定単語を避けた単語を出力する確率が高くなる。この結果、入力音声信号が、外部機器に悪影響を及ぼす単語あるいは外部機器を話者の意図と異なる状態で動作させる事態を招く単語として誤認識される可能性が低くなるから、その外部機器が不用意な動作をする事態を効果的に防止可能となる。

【0008】

請求項2記載の手段によれば、認識対象候補集団として、外部機器の機能低下を招く特定の単語と類似した音声群に対応した比較対象パターンの集団が用意されるから、入力音声信号が、外部機器の機能低下を招く単語として誤認識されることに起因して当該外部機器が不用意な動作をする事態を効果的に防止可能となる。

10

【0013】

前記目的を達成するために、請求項3に記載した手段を採用することができる。この手段においても、音響分析手段、特徴抽出手段及びパターンマッチング手段を備えており、話者による入力音声信号を音響分析すると共に、その分析結果に基づいて入力音声信号の特徴パラメータを抽出し、抽出された特徴パラメータと予め記憶されている認識対象語彙の比較対象パターンとのパターンマッチングを行って一致度合いが高い単語を認識結果とするものであり、その認識結果を外部機器の動作制御に供するようになる。この場合、認識対象語彙の比較対象パターンとして、外部機器を話者の意図と異なる状態で動作させる可能性のある特定の単語に対応した比較対象パターンの他に、その特定単語と類似した音声群に対応した比較対象パターンの集団が認識対象候補集団として用意されており、パターンマッチング手段は、パターンマッチング時において、入力音声信号から抽出された特徴パラメータとの一致度合いが高い比較対象パターンが上記認識対象候補集団中に存在する場合に、前記特定単語を避けた単語を出力する。

20

【0014】

ここで、認識対象語彙の比較対象パターン中に、外部機器を話者の意図と異なる状態で動作させる可能性のある特定単語に対応した比較対象パターンの他に、その特定単語と類似した音声群に対応した比較対象パターンの集団が用意された場合には、入力音声信号が特定単語であると認識される特徴空間が限定されることになる。このような状態は、特定単語以外の単語が当該特定単語と認識される可能性が相対的に低くなった状態に相当するようになり、このときには、当該特定単語を避けた単語を出力する確率が高くなる。この結果、入力音声信号が、外部機器に悪影響を及ぼす単語あるいは外部機器を話者の意図と異なる状態で動作させる事態を招く単語として誤認識される可能性が低くなるから、その外部機器が不用意な動作をする事態を効果的に防止可能となる。

30

【0016】

請求項4記載の手段によれば、認識対象候補集団として、外部機器の設定内容の変更を招く可能性のある特定の単語と類似した音声群に対応した比較対象パターンの集団が用意されるから、入力音声信号が、外部機器の設定内容の変更を招く単語として誤認識されることに起因して当該外部機器が不用意な動作をする事態を効果的に防止可能となる。

【0019】

請求項6に記載したように、パターンマッチング手段が、前記特定単語を避けた認識結果を出力する動作を、外部機器における所定の動作制御を許容するか否かの確認のために話者による音声入力を求める場面で行う構成とした場合には、外部機器の動作の有無を最終的に確認するための入力音声信号が誤認識される可能性が小さくなるため、その外部機器の不用意な動作を防止する上で有益となる。

40

【0020】

請求項7に記載した手段によれば、ナビゲーション装置のボイスコントロールが可能になると共に、入力音声信号の誤認識に起因した装置の誤動作を効果的に防止できるようになる。

【0021】

請求項8に記載したように、上記した音声認識装置におけるパターンマッチング手段の機

50

能は、その機能をコンピュータシステムにて実現するためプログラム及びデータを記録した記録媒体により得ることができる。この場合、例えば、CD-ROM、DVD-ROM、不揮発性メモリ（EEPROMなど）、フロッピー（R）ディスク、光磁気ディスク、ハードディスクなど、多様な記録媒体を利用でき、その記録媒体に記録されたプログラム及びデータを必要に応じてコンピュータシステム側に読み込み、そのプログラム及びデータに従った情報処理を実行することによりパターンマッチング手段の機能を実現できる。

【0023】

【発明の実施の形態】

（第1の実施の形態）

以下、本発明をカーナビゲーション装置に適用した第1実施例について図1ないし図4を参照しながら説明する。

図4には、カーナビゲーション装置1を含む全体の電氣的構成が概略的に示されている。この図4において、カーナビゲーション装置1は、位置検出器2、地図データ入力器3、操作スイッチ群4、これらに接続されたマイクロコンピュータを主体として成る制御回路5、この制御回路5に接続された外部メモリ6、例えばフルカラー液晶ディスプレイから成る表示装置7、リモコンセンサ8、対話制御部9、音声合成部10、例えばアンプ内蔵型のスピーカ11を備えた構成となっている。

【0024】

位置検出器2は、周知構成の地磁気センサ12、ジャイロセンサ13、距離センサ14、及びGPS衛星からの電波に基づいて車両の位置を検出するためのGPS受信機15を有しており、これら各センサ12～15は、車両の適宜部所に設置される。前記制御回路5は、位置検出器2の各センサ12～15が性質の異なる誤差を有しているため、各々補間しながら使用するように構成されており、これらセンサ12～15からの入力に基づいて、車両の現在位置、進行方向、速度や走行距離などを高精度で検出する。

【0025】

前記地図データ入力器3は、道路地図データや、位置検出精度の向上のための所謂マップマッチング用データなどを含む各種データを記憶した記憶媒体からデータを入力するためのドライブ装置から成り、その記憶媒体としては、例えばCD-ROMやDVD-ROMなどの大容量記憶媒体が用いられる。尚、前記道路地図データは、道路形状、道路幅、道路名、建造物、各種施設、それらの電話番号、地名、地形などのデータを含むと共に、その道路地図を前記表示装置7の表示画面上に再生するためのデータを含んで構成されている。

【0026】

前記操作スイッチ群4は、ユーザ（運転者）が、目的地の指定や、表示装置7に表示される道路地図の選択などの各種のコマンドを入力するための各種のメカニカルスイッチから構成されている。また、この操作スイッチ群4の一部は、前記表示装置7の画面上に設けられたタッチスイッチ（図示せず）からも構成されるようになっている。そして、操作スイッチ群4と同等の機能を有するリモートコントロール端末16（以下、リモコンと称する）も設けられており、このリモコン16からの操作信号が、前記リモコンセンサ8により検出されるようになっている。

【0027】

前記表示装置7の画面には、各種縮尺の道路地図が表示されると共に、その表示に重ね合わせて、車両の現在位置及び進行方向を示すポイントが表示されるようになっている。また、ユーザが目的地などを入力するための各種の入力用画面や、各種のメッセージやインフォメーションなども表示されるようになっている。さらには、目的地までの案内を行うルートガイダンス機能の実行時には、道路地図に重ね合わせて進むべき経路などが表示されるようになっている。

【0028】

そして、前記制御回路5は、地図データ入力器3からの道路地図データに基づいて表示装置7に道路地図を表示させると共に、位置検出器2の検出出力に基づいて車両の現在位置

及び進行方向を示すポインタを表示させるようになっている。このとき、車両の現在位置を道路地図上に表示するアップマッチングが行われるようになっている。また、ユーザのコマンド入力に基づいて、表示装置 7 に表示させる地図の種類（縮尺）の切替えなどを行うようになっている。

【0029】

さらに、制御回路 5 は、ユーザによる目的地の入力に基づいて、自動ルート探索及びルートガイダンスの機能を実行するようになっている。詳しい説明は省略するが、自動ルート探索の機能は、車両の現在位置からユーザにより入力された目的地までの推奨する走行経路を自動的に算出するものであり、ルートガイダンスの機能は、上述のように、表示装置 7 の画面にその走行経路を表示して目的地まで案内するものであり、このとき、対話制御部 9 を通じて音声合成部 10 の音声合成機能を制御することにより、例えば「200m先の交差点を左です」といった音声スピーカ 11 から出力させる音声案内も併せて行うことができるようになっている。

【0030】

上記対話制御部 9 には、マイクロホン 17 を介して入力された音声信号を認識するための音声認識装置 18 からの認識結果が与えられるようになっている。この音声認識装置 18 は、カーナビゲーション装置 1 に対するコマンド（周知の行き先検索指示、リルート指示、表示地図の縮尺変更指示など）の入力や目的地（自宅、登録ポイントなど）の指示などを、前記操作スイッチ群 4 あるいはリモコン 16 の手動操作に代えて、ユーザ（運転者）が前方を見たまま音声入力することによって同様に行い得るようにしたものであり、これにより、安全性、利便性を向上させるための装置として設けられている。尚、図示しないが、マイクロホン 17 は、車両の例えばステアリングコラムカバーの上面部や運転席側のサンバイザーなどの運転者の声を拾いやすい位置に設けられるものである。

【0031】

図 1 には、上記音声認識装置 18 の概略構成が機能ブロックの組合わせにより示されており、以下これについて説明する。

図 1 において、音声認識装置 18 は、音響分析部 19 と、特徴抽出部 20 と、パターンマッチング部 21 とを備えた構成となっている。尚、これら音響分析部 19、特徴抽出部 20、パターンマッチング部 21 は、それぞれ本発明でいう音響分析手段、特徴抽出手段、パターンマッチング手段に相当する。

【0032】

音響分析部 19 は、マイクロホン 17 を介して入力されたアナログ音声信号を、例えば 12 kHz 程度のサンプリング周期でデジタル信号に変換し、その変換信号をオーバーラップさせながら一定時間毎に順次所定長さのフレーム信号として切り出す。そして、フレーム毎のデジタル音声入力信号にフーリエ変換を行うことによって、入力音声信号の音声スペクトルを求め、さらに、その音声スペクトルの振幅の 2 乗を計算して短時間音声パワースペクトルを得る。そして、このように得たパワースペクトルに対し逆フーリエ変換を施して自己相関係数を求め、この自己相関係数を用いた LPC 分析を行って LPC 係数を計算するものであり、その LPC 係数は特徴抽出部 20 へ順次送られる。

【0033】

特徴抽出部 20 は、音響分析部 19 から送られてきた LPC 係数に基づいて、フレーム毎のスペクトル上の特徴パラメータとしての LPC ケプストラム係数を計算するものであり、この LPC ケプストラム係数はパターンマッチング部 21 へ順次送られる。

【0034】

パターンマッチング部 21 は、辞書部に予め計算して記憶しておいた認識対象語彙の標準パターン（特徴パラメータ系列：本発明でいう比較対象パターンに相当）と、特徴抽出部 20 から送られてきた LPC ケプストラム係数の時系列データとの間でパターンマッチング（類似度計算）を行う。この場合、音声認識のアルゴリズムとして周知の HMM（隠れマルコフモデル）などによって、時系列データをいくつかの区間に分け、各区間が辞書データとして格納されているどの音声に類似しているかを、例えば尤度比に基づいて求める。

【0035】

ここで、パターンマッチング部21は、基本的には、認識対象語彙のうち、上記のような計算により得られた類似度が最も高い単語（または類似度が上位の複数の単語）を認識結果として決定し、その認識結果を対話制御部9へ出力する動作を行うものであるが、本実施例では、認識対象語彙の標準パターン（比較対象パターン）として以下に述べるような「認識対象候補集団」を用意した構成に大きな特徴を有するものである。

【0036】

即ち、上記認識対象候補集団としては、ボイスコマンドとしてカーナビゲーション装置1に与えられたときに、そのカーナビゲーション装置1の動作に機能低下などの悪影響を及ぼす可能性や、カーナビゲーション装置1がユーザ（話者）の意図と異なる状態で動作する可能性がある特定の単語（つまり、非常に高い認識率が要求される単語）と類似した音声群に対応した標準パターンの集団を含んだものが用意され、また、必要に応じて、上記特定単語と類似した音声群に対応した標準パターンの集団に加えて、当該特定単語と反対の意味を持つ第2の特定単語と類似した音声群に対応した標準パターンの集団も含んだものが用意される。尚、ここでいう「類似」とは「似ているが異なる」という意味である。

【0037】

要するに、上記特定単語としては、カーナビゲーション装置1が当該特定単語に対応したボイスコマンドを受けたときに取り返しのつかない動作（例えば、ルート探索動作により設定された目的地までのルートの破棄を含む動作）を実行する単語が選定される。また、上記第2の特定単語と類似した音声群に対応した標準パターンの集団も含んだ認識対象候補集団が必要となる具体的な例としては、以下のような事例が考えられる。

【0038】

即ち、カーナビゲーション装置1において、目的地までの設定ルートの変更を伴うボイスコマンド（例えば、「リルート」、「自宅」、「家に帰る」、「行き先変更」など）が入力された場合には、ユーザに対して、設定ルートを変更する制御（つまり既設定ルートの破棄）を本当に行っても良いか否かを確認するために、ユーザによる音声入力を求める音声ガイダンスを発するものであり、この音声ガイダンスに対して、ユーザー側では「はい」／「いいえ」、「YES」／「NO」などのような二者択一のボイスコマンドを返答することになる。この場合の「はい」あるいは「YES」が上記特定単語に相当し、「いいえ」あるいは「NO」が上記第2の特定単語に相当するものであり、このような二者択一態様のボイスコマンドが入力される場合には、第2の特定単語と類似した音声群に対応した標準パターンの集団も含んだ状態の認識対象候補集団が必要になる。

【0039】

そして、この場合において、パターンマッチング部21は、認識結果の候補中に認識対象候補集団中の標準パターンに対応した音声が存在する場合に前記第2の特定単語を認識結果として出力する構成となっている。

【0040】

具体的には、例えば、特定単語が「YES」であった場合には、これと反対の意味を持つ「NO」が第2の特定単語とされるものであり、前記認識対象候補集団には、「YES」及び「NO」と類似した例えば、/j//*///s/、/j//e//*/、/au/、/uu/、……（但し、* は任意の母音あるいは子音が当てはまるという意味）という音声群に対応した標準パターンを含めた状態とされる。

【0041】

従って、この例のように、「YES」という特定単語について上記のような「NO」と類似した音声群に対応した標準パターンの集団を含めた認識対象候補集団が用意された場合には、パターンマッチング部21でのパターンマッチング時において、「YES」と反対の意味の「NO」に多くの言い換えが用意された状態と等価となって、ユーザによる入力音声信号が「YES」であると認識される特徴空間が図2(a)に示すように限定されることになる（○印が「YES」に対応した標準パターン、×印が「NO」に対応した標準

10

20

30

40

50

パターン並びに認識対象候補集団中の各標準パターンに相当)。その結果、入力音声信号が「YES」であるとの確信が低い場合に、これと反対の意味を持つ「NO」を認識結果として出力する確率が高くなり、これによりカーナビゲーション装置1がユーザにとって不用意な動作をする事態を効果的に防止可能となる。

【0042】

因みに、従来構成では、ユーザによる入力音声信号が「YES」であると認識される特徴空間が図2(b)に示すように広くなるため、例えば、「YES」以外の音声を発生したときに「YES」と誤認識される率が高くなるものである。このような事象を確認するために、本件発明者は、図1に示す基本構成を有する音声認識装置18において、実際に「YES」、「NO」と発生したときに「YES」、「NO」の認識結果がそれぞれ出力される率、並びに「YES」、「NO」以外の音声を発生したときに「YES」の認識結果が出力される率を、認識対象語彙の標準パターンとして前述した認識対象候補集団を用意しない場合(従来構成に相当)と用意した場合とについて実験により求めた。その実験結果をまとめた図3に示すように、認識対象候補集団を用意していない『対策前』の認識率と、当該認識対象候補集団を用意した『対策後』の認識率との間には大きな差が出ることが判明した。尚、図3中には、実験サンプル数と該当出力数との関係も示した。要するに、この図3からは、本実施例の構成によれば、「NO」の認識率が向上すること、「YES」及び「NO」と無関係の音声を「NO」と認識する確率が上がること、「YES」の認識率には影響しないことが分かる。

【0043】

尚、ここでは、前記認識対象候補集団中に、第2の特定単語(「NO」が一例)と類似した音声群に対応した標準パターンの集団も含めた構成について言及したが、これらの標準パターンの集団は、認識対象候補集団中に必要に応じて含める構成とすれば良いものである(少なくとも、特定単語と類似した音声群に対応した標準パターンの集団を認識対象候補集団中に含める構成とすれば良い)。

【0044】

以上要するに、上記した本実施例の構成によれば、カーナビゲーション装置1側において設定ルートを変更する制御を本当に行っても良いか否かを確認するために、ユーザによる音声入力を求める音声ガイダンスを発する場面において、ユーザによる入力音声信号が設定ルートの変更を許容する特定単語(実施例では「YES」)であるとの確信が低い場合に、当該特定単語と反対に意味を持つ第2の特定単語(実施例では「NO」)を出力する確率が高くなり、これによりカーナビゲーション装置1が、ユーザの意図に反した不用意な動作をする事態を効果的に防止可能になるものである。

【0045】

(第2の実施の形態)

図5には本発明の第2実施例が示されており、以下これについて前記第1実施例と異なる部分のみ説明する。

【0046】

この第2実施例では、前記第1実施例におけるパターンマッチング部21とは異なる処理動作を行うパターンマッチング部21'(パターンマッチング手段に相当)を設け、このパターンマッチング部21'でのパターンマッチング動作に供するための認識対象語彙の標準パターンに対応した辞書データとして、第1実施例における認識対象候補集団を用意することなく、一般的な辞書データを用意している。また、本実施例では、音響分析部19、特徴抽出部20あるいはパターンマッチング部21'の少なくとも何れかにおいて得られる所定の信頼度評価用特徴量に基づき、パターンマッチング部21'による絶対的な信頼度を推定する信頼度推定手段を設ける構成としている。

【0047】

具体的には、例えば図5に示すように、音声認識装置18中に信頼度推定手段として正解確率推定部22が設けられる。この正解確率推定部22は、パターンマッチング部21'にて得られた認識結果の信頼度を適切に評価可能な指標として、その認識結果の正解確率

10

20

30

40

50

を推定し、当該パターンマッチング部21'にフィードバックする構成となっている。上記正解確率推定部22は、例えば、パターンマッチング部21'から「信頼度評価用特徴量」として得られる単語と音節連接モデルの尤度比や単語内の音節の継続時間の分散に基づいて認識結果の正解確率の推定を行うもので、詳しくは、本発明の出願人と同一人の出願に関わる特願平11-359076号（発明の名称：音声認識装置及び記録媒体）に記載された技術を使用する。

【0048】

尚、ここでは、信頼度評価用特徴量をパターンマッチング部21から得る構成としたが、図5に破線で示すように、音響分析部19や特徴抽出部20から得るようにしても良い。その場合、音響分析部19から得る「信頼度評価用特徴量」としては、音響分析自体の成功度合いを示す特徴量が考えられ、例えばノイズ度合いやLPC残差が挙げられる。また、特徴抽出部20にて得られる「信頼度評価用特徴量」としてはケプストラム係数が考えられる。

10

【0049】

パターンマッチング部21'は、パターンマッチング時において、認識結果の候補中に前記第1実施例で述べたような特定単語（非常に高い認識率が要求される単語）が含まれる場合に、その特定単語の認識結果の信頼度を示す正確確率を正確確率推定部22から取り込むと共に、その正確確率が所定レベル（このレベルは実験的に求めることができる）以下であった場合に、当該特定単語の認識信頼性が低いと判断し、その特定単語を避けた他の単語を認識結果として出力する構成となっている。

20

【0050】

要するに、パターンマッチング部21'は、辞書部に予め計算して記憶しておいた認識対象語彙の標準パターンと、特徴抽出部20から送られてきたLPCケプストラム係数の時系列データとの間でのパターンマッチング時において、認識結果の候補となった単語の認識信頼性が絶対的に低く、且つその候補単語が特定の単語（カーナビゲーション装置1の動作に機能低下などの悪影響を及ぼしたり、カーナビゲーション装置1が設定内容の不用意に変更などユーザの意図と異なる状態で動作する可能性がある単語）であった場合に、その特定単語を避けた単語を出力する動作を行うものである。

【0051】

この結果、本実施例によれば、入力音声が入力音声装置1の動作不都合状態（機能低下や設定内容の変更）を招く可能性がある特定の単語として誤認識される可能性が低くなるから、カーナビゲーション装置1が不用意な動作をする事態を未然に防止できるようになる。また、この場合には、パターンマッチング部21'での認識結果の絶対的な信頼度を、その認識結果が正解である確率に基づいて推定する構成となっているから、その認識結果の信頼度を直観的且つ定量的に把握できるようになり、当該認識結果を利用するアプリケーションを実際に構築する際に便利になる。

30

【0052】

尚、本実施例において、認識結果の絶対的な信頼度を推定する正確確率推定部22は以下のような事情に対処して設けられるものである。即ち、パターンマッチングにより得られる尤度（確からしさ）は、標準パターン（比較対象パターン）との一致度合いを示すものであって、あくまでも相対的なものである。つまり、パターンマッチング時において尤度が一番大きなものを認識結果として選択したとしても、それは相対的に一致度合いが高いだけであり、実際の入力音声信号に対して、その認識結果がどの程度正しいかという判断に関しては適切な評価ができない。このような事情に対処するために、本実施例では、パターンマッチング手段21'での認識結果を絶対的な観点で評価可能な「信頼度」という指標を、音響分析部19、特徴抽出部20あるいはパターンマッチング部21'の少なくとも何れかにおいて得られる所定の「信頼度評価用特徴量」に基づいて推定するようにしたものである。

40

【0053】

（その他の実施の形態）

50

尚、本発明は上記した各実施例に限定されるものではなく、次のような変形または拡張が可能である。

第1実施例におけるパターンマッチング部21の機能をコンピュータシステムにて実現するために必要な処理プログラム並びに認識対象語彙の標準パターンや認識対象候補集団などのデータを記録した記録媒体を、CD-ROM、DVD-ROM、不揮発性メモリ（EEPROMなど）、フロッピーディスク、光磁気ディスク、ハードディスクなど、多様な記録媒体を利用して提供することができる。さらに、第2実施例におけるパターンマッチング部21'及び正解確率推定部22の機能をコンピュータシステムにて実現するために必要な処理プログラム並びに認識対象語彙の標準パターンやその他の必要なデータを記録した記録媒体も同様に提供することもできる。第2実施例では、パターンマッチング部21'による認識結果の絶対的な信頼度を推定する信頼度推定手段として正解確率推定部22を用いる構成としたが、これ以外の手段を採用することも可能である。さらに、動作制御対象の外部機器として、個人が携帯するモバイル型ナビゲーション装置、あるいはその他の一般的な機器を選定することができる。

【図面の簡単な説明】

【図1】本発明の第1実施例を示す音声認識装置の概略的な機能ブロック図

【図2】音声信号を認識をする特徴空間の例を模式的に示す図

【図3】実験結果を示す図

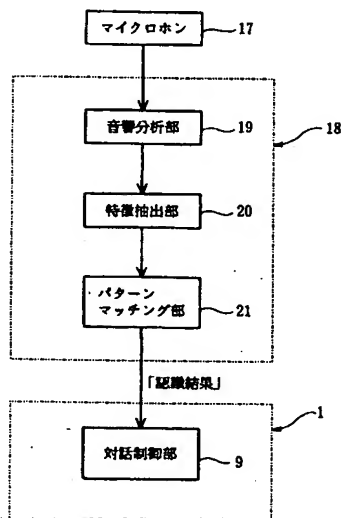
【図4】全体の電氣的構成を概略的に示す機能ブロック図

【図5】本発明の第2実施例を示す図1相当図

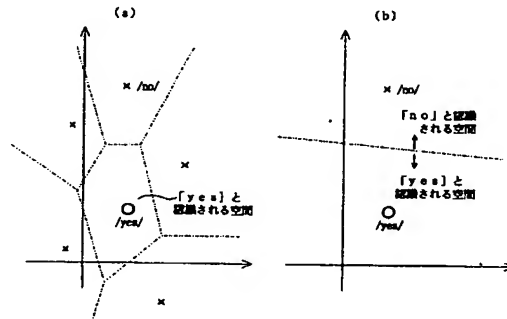
【符号の説明】

1はカーナビゲーション装置、9は対話制御部、17はマイクロホン、18は音声認識装置、19は音響分析部（音響分析手段）、20は特徴抽出部（特徴抽出手段）、21、21'はパターンマッチング部（パターンマッチング手段）、22は正解確率推定部を示す。

【図1】



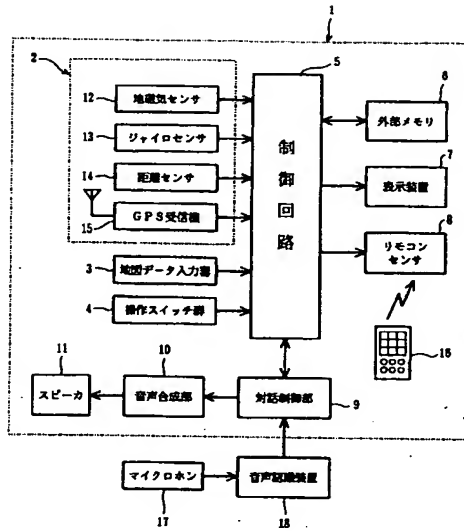
【図2】



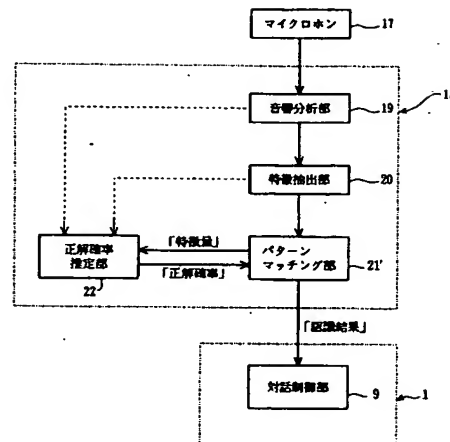
【図3】

	「yes」の認識率 (出力数/サンプル数)	「no」の認識率 (出力数/サンプル数)	その他から「yes」に 認識した率 (出力数/サンプル数)
対策前	100(18/18)	95.3(485/510)	61.9(915/1477)
対策後	100(18/18)	98.2(501/510)	8.7(128/1477)

【図 4】



【図 5】



フロントページの続き

審査官 渡邊 聡

(56)参考文献 特開2000-069348 (JP, A)
特開平07-181994 (JP, A)
特開平09-244691 (JP, A)
特開平08-211892 (JP, A)
特開平11-337365 (JP, A)
特開2001-222291 (JP, A)
特開2002-021121 (JP, A)
特開2000-069348 (JP, A)
特開平07-181994 (JP, A)
特開平09-244691 (JP, A)
特開平08-211892 (JP, A)
特開平11-337365 (JP, A)

(58)調査した分野(Int.Cl.⁷, DB名)

G10L 15/28

G10L 15/06